Incorporating Psychophysical Models in Image Synthesis

Brian Guenter
College of Computing
Georgia Institute of Technology
Atlanta GA 30332-0280

email: brian.guenter@cc.gatech.edu

1. Psychophysical Models in Image Synthesis

It has long been known that the human visual system (HVS) is not a camera; the eye-retina-brain combination does not simply record a visual scene the way a movie camera or video camera does. Instead the HVS processes the image in a series of steps of increasing abstraction. At each step some information about the scene may be lost and replaced by a higher level representation. Thus the HVS is remarkably insensitive to the loss of certain kinds of information in an image scene.

Conversely, since the HVS has evolved to detect meaningful patterns, it is exceptionally sensitive to certain image characteristics, a case in point being vernier acuity. Vernier acuity refers to the ability of the HVS to estimate relative displacements between lines or other image features to a resolution much finer than the spacing between the rods or cones in the fovea . Vernier hyperacuity accounts for the exceptional visibility of “jaggies”, the ragged line appearance of lines drawn on a computer display without proper antialiasing .

The desirability of taking the characteristics of the HVS into account when designing image synthesis algorithms has been pointed out before . However, very little research on the application of visual psychophysical results in image synthesis algorithms has been done. However extensive investigation has been carried out in connection with reduced bandwidth HDTV systems and improved definition systems compatible with the NTSC standard. Most of the research of interest has been applied to reducing transmission bandwidth without significantly reducing image quality, or to improving the quality of signals compatible with the current NTSC standard.

These bandwidth reduction and image improvement techniques are not directly applicable to the image synthesis problem, partially because some of them are concerned with overcoming the weaknesses of the NTSC encoding system, but primarily because they assume as input a high quality image sequence at a given spatial and temporal resolution which is then processed to remove relatively imperceptible information. This is true of image compression techniques in general, not just those of interest in the broadcast television industry.

In image synthesis however, one wants to predict what will eventually be imperceptible and then avoid computing it. Consequently techniques which work quite well for reducing transmission bandwidth or image storage size may be completely useless for reducing computation in image synthesis.

While image compression and bandwidth reduction techniques cannot be directly applied to the image synthesis problem they can suggest which psychophysical properties of the HVS to exploit. Any HVS characteristic which has been exploited to yield a factor of two or so reduction in bandwidth with little or no perceptible loss of image quality is a good candidate for use in reducing computation in image synthesis. It may be the case, however, that the method used in image synthesis might be very different from that used in bandwidth reduction even though the HVS characteristic being exploited might be identical.

2. Incorporating Psychophysical Models in Image Synthesis

In the past filtering in image synthesis algorithms has been used almost exclusively for antialiasing and the antialiasing filters have not been designed with the characteristics of the HVS in mind. It is possible to do much more with filtering by capitalizing on the properties of the HVS. The general idea of incorporating psychophysical models into image synthesis is to:

- choose a characteristic of the HVS to exploit
- design a filter response to match, enhance, or filter out the image characteristic
- implement the filter in an efficient way

The first step requires a knowledge of the spectral characteristics of the HVS. Many psychophysical measurements are taken at or near threshold where the response of the HVS is assumed to be linear over a small range. Most viewing occurs under conditions far from threshold. In spite of the fact that the linearity assumption is a gross approximation algorithms based on this assumption can give useful results. The linearity assumption makes it possible to use linear filtering which is much easier to analyze and implement.

With the spectral characteristic in hand the second step requires designing a two or three dimensional filter. This filter will perform one of three functions. It may match the response of the image synthesis algorithm to that of the HVS characteristic, thereby avoiding the computation of portions of the spectrum which will be imperceptible. It may enhance the response of the image synthesis algorithm so that the particular characteristic becomes more visible - this is the case with edge sharpening filters for example. It may filter out the image characteristic - this is the case with antialiasing filters, which eliminate edge jagginess.

The third step requires developing an algorithm that is more efficient than an image synthesis algorithm which does not incorporate a psychophysical model. This requires some ingenuity and an extension of the conventional image synthesis pipeline and is the subject of the next section.

2.1. More General Filtering in Image Synthesis

Let us examine how previous image synthesis algorithms have performed filtering and sampling. A typical image synthesis filtering pipeline for periodic sampling is shown in Figure 1. The image is either prefiltered and then sampled on a rectangular lattice or is sampled and then filtered with a discrete filter, again on a rectangular lattice. The rectangular lattice used in the image synthesis algorithm matches that of the display device so no sample lattice conversion is necessary.

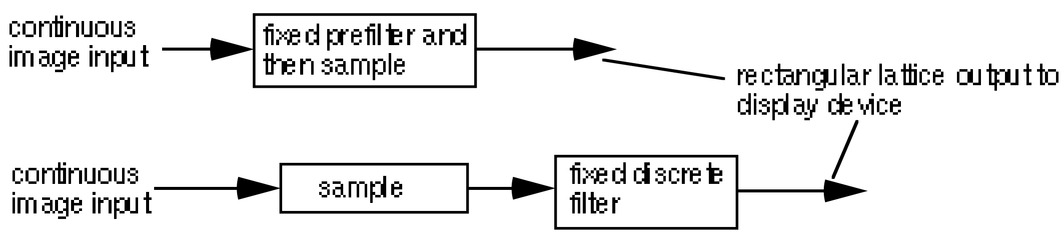

Figure 1 Conventional image synthesis filtering pipeline

The new image synthesis filtering pipeline is more flexible. There can be as many as three filtering stages: a *prefilter* stage, a *discrete filter* stage, and a *sample lattice conversion* stage. Each stage accepts both filter and sampling lattice specifications. The filters and sampling lattices may be two or three dimensional. The filter specification for the *prefilter* stage will be for a continuous filter; the filter specifications for the *discrete filter* and *sample lattice conversion* stages will both be discrete.

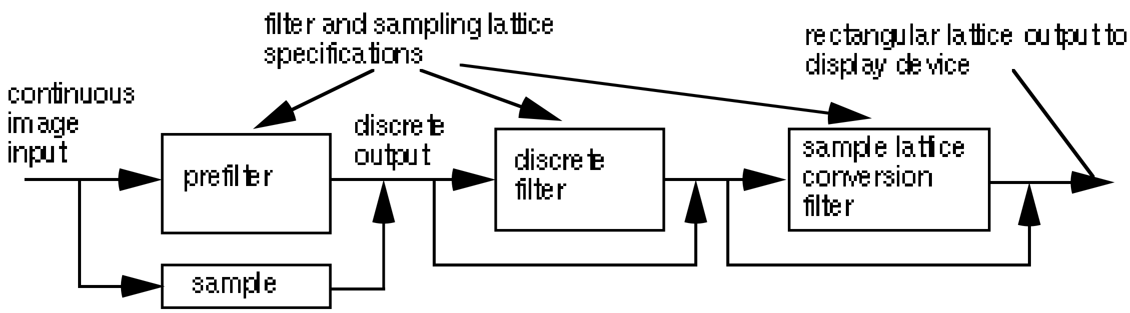

Figure 2 New image synthesis filtering pipeline

The *prefilter stage* filters the image in the continuous domain and then performs sampling. Several two and three dimensional prefilters have been developed . These filters tend to be complex and computationally expensive. In addition some will only handle radially symmetric two dimensional impulse responses . Others will handle anisotropic impulse responses but at the expense of a three dimensional look up table, which can be unreasonably large. For general purpose spectrum shaping it is necessary to have a simple, efficient prefilter which can handle anisotropic impulse responses without requiring unreasonable amounts of space or computation. A spatial prefilter requiring only a two dimensional look up table and relatively low computation has been developed . It is being used to explore the application of anisotropic spatial prefiltering to exploit the psychophysical properties of the HVS.

The *discrete filter stage* performs filtering in the discrete domain. The discrete filter stage may be used instead of, or together with, the prefilter stage. Discrete filtering is a well understood operation and extensions to non-rectangular sampling lattices are straightforward .

The *sampling lattice conversion stage* is used to convert from the optimal sampling lattice used in the previous two filtering stages to the sampling lattice used by the display device. Sampling lattice conversion may involve either or both of two operations: decimation and interpolation. Interpolation increases the sample rate and decimation decreases the sample rate. To avoid the introduction of aliasing artifacts filtering must be performed for both interpolation and decimation.

Any or all of these stages may be bypassed. For example, it would be possible to bypass the prefilter and simply sample the image on a rectangular sampling lattice, filter in the discrete filter stage, and skip the sample lattice conversion filter since the output is already in the rectangular display lattice format.

In practice the discrete filter and sample lattice conversion stages may be combined into a single filtering operation but it is important to separate them conceptually because they perform different functions. The discrete filter shapes the frequency response to take advantage of some psychophysical characteristic of the HVS; the sample lattice conversion filter merely converts from the sampling lattice of the discrete filter to that of the display device.

Sample lattice conversion in two dimensions is relatively straightforward if linear spatially invariant filtering is used but it is well known that linear spatially invariant filters have unavoidable tradeoffs between apparent image sharpness, ringing, and rejection of spectral replicas . The best rejection of spectral replicas occurs when an ideal low pass filter is used but this filter has intolerable ringing. Filters which reject spectral replicas and have good apparent sharpness typically have negative lobes. This causes black level clipping, an objectionable image artifact . Non-linear or linear spatially variant filters may be able to overcome these tradeoffs.

In three dimensions the lattice conversion problem is considerably more complicated since linear spatially invariant interpolation filters cause multiple images when there is motion in the image. Here spatially variant, and possibly non-linear, filtering is a requirement for acceptable results.

With this more flexible pipeline filtering can be performed entirely by the prefilter, entirely by the discrete filter or by both. Dual domain filtering may prove useful because a prefilter capable of handling arbitrary filter impulse responses is computationally expensive and the algorithms are complex enough that they are difficult to put into hardware. Discrete filters are easy to put into hardware or program into current high speed digital signal processing chips but they can never completely eliminate aliasing artifacts because they are applied after sampling has occurred.

A combination of a prefilter with a complementary discrete filter could provide the superior antialiasing characteristics of a prefilter with the simplicity of a discrete filter. Certain specialized prefilters, such as the two dimensional box filter, are relatively simple to program and fast. Then, for example, a box prefilter could be combined with a complementary discrete filter to yield the desired composite response. This kind of hybrid solution may well provide superior antialiasing and frequency shaping characteristics at lower computational cost than either prefiltering or discrete filtering alone.

It is important to understand that the frequency response and sampling lattice have to be designed together to achieve the lowest sample rate. In the one dimensional case the sampling rate determines how far apart the spectral replicas of the baseband signal will be in frequency space. In the multidimensional case the situation is more involved. Certain sampling lattices are more efficient than others for given spectral shapes, i.e., the sampling rate needed by these optimal lattices to exactly represent multidimensional bandlimited signals is the lowest possible.

Unfortunately, optimal sampling lattices are not usually rectangular although display sampling lattices are virtually always rectangular. If a signal is sampled on an optimal lattice then it must be converted to a rectangular sampling lattice for display. This conversion process may involve one or both of two operations: interpolation and decimation. Interpolation increases the sampling rate of the signal and decimation reduces the sampling rate. See for a more detailed explanation of sample rate conversion in one dimension, and for an extension of these techniques to multidimensional signals.

Choosing an optimal sampling lattice reduces to finding the closest packing for a given spectral shape. For hyperspherically bandlimited signals the closest packing lattice has been determined for dimensions up to 8 . As the dimension of the signal increases cubical sampling becomes increasingly inefficient for hyperspherically bandlimited signals: in three dimensions only .705 as many samples are needed for the closest packing lattice compared to cubical sampling; in four dimensions the ratio drops to .499.

Three and four dimensions are particularly interesting for computer graphics purposes because they are so common. Computer animation, which consists of a sequence of images, is a three dimensional signal. Examples of four dimensional signals would be image sequences of fluid flow or electric field intensity. Such image sequences are becoming increasingly common because of the increased interest in scientific visualization. For these high dimensional signals the savings from using an optimal sampling lattice can be substantial but interpolation between the optimal lattice and the rectangular display lattice becomes more complex as the dimensionality of the signal increases.

2.2. An Example

A concrete example should make clear how the ideas discussed so far can be applied to a real problem. Let us begin by assuming that we would like to take advantage of the anisotropic spatial response of the HVS to reduce computation and image storage requirements. The diagonal resolving power of the HVS is approximately .7 that of the horizontal and vertical resolving power . This means that if an image is filtered so that the highest diagonal frequencies are only .7 that of the highest horizontal and vertical frequencies the effect of the filter will be almost unnoticeable.

The first step in exploiting this effect is to design a filter frequency response shape which performs this diagonal filtering; the second step is to choose a sampling lattice which minimizes the sampling rate required for this filter. To emphasize the importance of using both frequency shaping and non-rectangular sampling together the effect of just frequency shaping with rectangular sampling will be contrasted with the effect of frequency shaping and optimal sampling together.

The desired filter should have a frequency cutoff along the diagonal at approximately .7 that along the horizontal and vertical and should be shaped so that the spectral replicas can be packed very tightly. Such a frequency response is shown in Figure 3. Simple geometry shows that the highest diagonal frequency is approximately .707 that of the highest horizontal or vertical frequency. This spectral shape also has the advantage that it can be packed as tightly as possible since the plane can be tiled with replicas of this baseband spectrum.

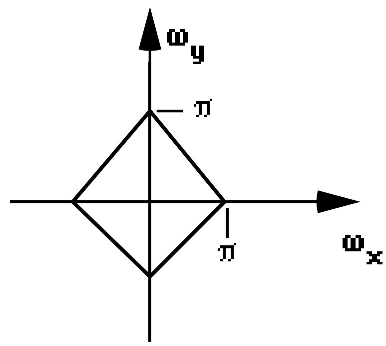

Figure 3 Frequency response of a diagonal low pass filter

Let us assume that an image has been diagonally low pass filtered and is sampled on a square lattice with sampling matrix

$$\mathbf{V}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{1}$$

This sampling lattice will have the basis vectors shown in Figure 4.

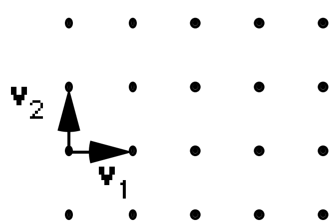

Figure 4 Basis vectors for a square sampling lattice

From (eq. 12) the corresponding spectral periodicity matrix will be

$$\mathbf{U}_1 = \begin{bmatrix} 2\pi & 0 \\ 0 & 2\pi \end{bmatrix} \tag{2}$$

and the baseband spectrum will be replicated as shown in Figure 5.

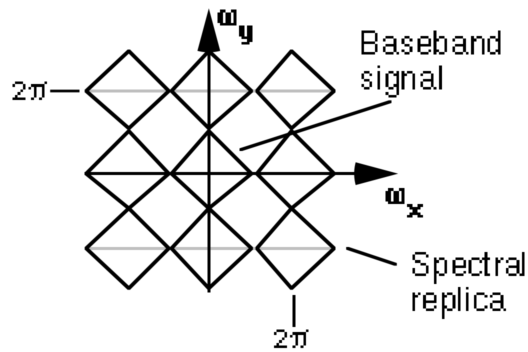

Figure 5 Spectral replication for square sampling lattice

Notice that even though there are large gaps where no signal energy is present the sampling rate cannot be reduced without introducing overlap, i.e. aliasing, between the spectral replicas. If rectangular sampling is used then no reduction in sampling rate is possible and nothing is gained by diagonal bandlimiting.

Now assume the same signal is sampled on a hexagonal lattice with sampling matrix

$$\mathbf{V}_2 = \begin{bmatrix} .5 & .5 \\ 1 & -1 \end{bmatrix} \tag{3}$$

The sampling lattice for this sampling matrix is shown in Figure 6 .

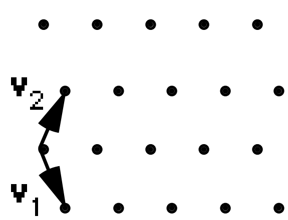

Figure 6 Hexagonal sampling lattice

The spectral periodicity matrix is

$$\mathbf{U}_2 = \begin{bmatrix} 2\pi & 2\pi \\ \pi & -\pi \end{bmatrix} \tag{4}$$

The baseband is replicated as shown on the left in Figure 7. Notice that now the sampling rate can be reduced by a factor of two in the x dimension without introducing overlap between the spectral replicas. So with the combination of spectral shaping and non-rectangular sampling it is possible to reduce the sample rate by a factor of two and still retain very high image quality. This technique has been used in reducing bandwidth requirements for HDTV and improved NTSC signals .

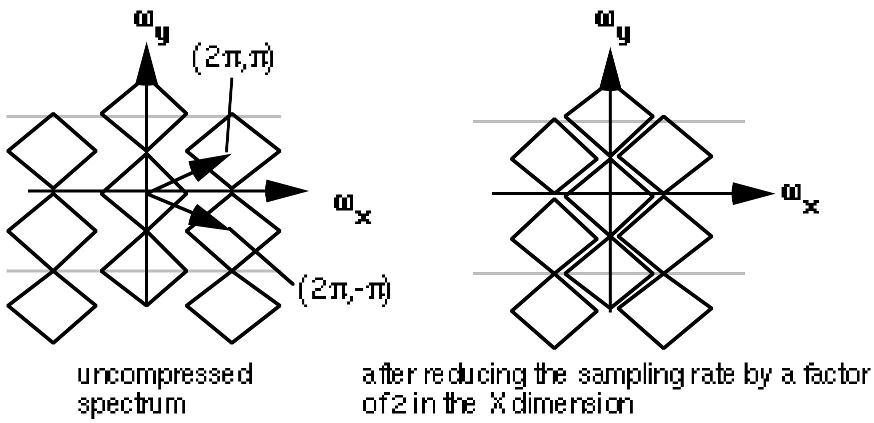

Figure 7 Spectral replication with hexagonal sampling

This example demonstrates all of the essential steps required in the new image synthesis paradigm:

• choose a characteristic of the HVS to exploit
• design a frequency response to match, enhance, or filter out the image characteristic
• choose an optimal sampling lattice for the frequency response chosen above
• design a filter with the required frequency response
• filter and sample the image using the optimal sampling lattice
• interpolate the image back to a rectangular display lattice

This application involved only two dimensional filtering but the technique generalizes to any number of dimensions.

Appendix: Sampling Lattices

Periodic sampling in N dimensions is slightly more complex than in the one dimensional case because there are many ways to periodically sample an N dimensional signal. A periodic sampling lattice can be generated by defining a set of linearly independent real vectors $\mathbf{v}_1,\mathbf{v}_2,...\mathbf{v}_N$. Points in the sampling lattice are generated by

$$\mathbf{t} = \mathbf{V}\mathbf{n} \qquad (5)$$

where $\mathbf{V}$, the periodicity matrix, has the lattice basis vectors as it columns

$$\mathbf{V} = \left[\mathbf{v}_1\ \mathbf{v}_2 ... \mathbf{v}_N\right] \qquad (6)$$

and $\mathbf{n}$ is a vector of integers. Figure 8 shows a lattice generated by the two vectors $\mathbf{v}_1=[2,0]^T$, $\mathbf{v}_2=[1,1]^T$.

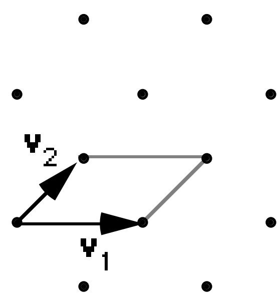

Figure 8

The corresponding periodicity matrix is

$$\mathbf{V} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \qquad (7)$$

The quantity $|\det \mathbf{V}|$ is equal to the volume of the hypercube with edges $\mathbf{v}_1...\mathbf{v}_N$. Physically this corresponds to the reciprocal of the sampling density. The continuous time signal $\mathbf{x}(\mathbf{t})$ has the Fourier transform

$$\mathbf{X}(\Omega) = \int_{-\infty}^{\infty} \mathbf{x}(\mathbf{t}) \exp(-j\Omega^T \mathbf{t})\, d\mathbf{t} \qquad (8)$$

If the signal $\mathbf{x}(\mathbf{t})$ is sampled then the sampled signal $\mathbf{y}(\mathbf{n}) = \mathbf{x}(\mathbf{V}\mathbf{n})$ has Fourier transform

$$\mathbf{Y}(\omega) = \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k}} \mathbf{X}(\mathbf{V}^{T^{-1}}(\omega - 2\pi\,\mathbf{k})) \tag{9}$$

or equivalently

$$\mathbf{Y}(\mathbf{V}^T\Omega) = \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k}} \mathbf{X}(\Omega - \mathbf{U}\mathbf{k}) \tag{10}$$

where **U**, known as the polar lattice matrix in the theory of the geometry of numbers , satisfies the equation

$$\mathbf{U}^T\mathbf{V} = 2\pi\mathbf{I} \tag{12}$$

The spectrum of the sampled signal consists of replicas of the continuous signal spaced in the frequency plane with periodicity matrix **U. This** is easiest to visualize in two dimensions. Suppose a signal with a diamond shaped passband is sampled with a rectangular lattice with sampling matrix

$$\mathbf{V}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{13}$$

The basis vectors for this sampling matrix and the associated sampling lattice are:

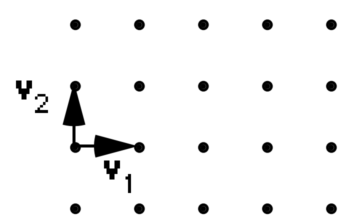

Figure 9

the corresponding spectral periodicity matrix is

$$\mathbf{U}_1 = \begin{bmatrix} 2\pi & 0 \\ 0 & 2\pi \end{bmatrix} \tag{14}$$

and the spectral lattice is:

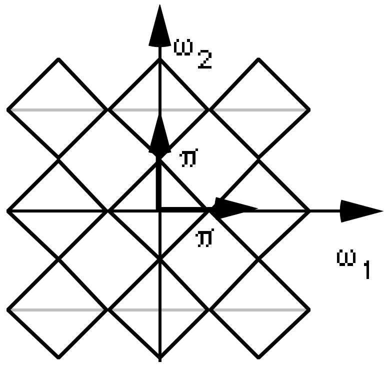

Figure 10

Now assume the same signal is sampled with a hexagonal lattice. The sampling matrix is

$$\mathbf{V}_2 = \begin{bmatrix} .5 & .5 \\ 1 & -1 \end{bmatrix} \tag{15}$$

and the lattice is:

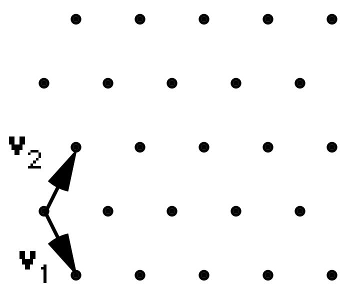

Figure 11

The spectral periodicity matrix is

$$\mathbf{U}_2 = \begin{bmatrix} 2\pi & 2\pi \\ \pi & -\pi \end{bmatrix} \tag{16}$$

and the lattice is:

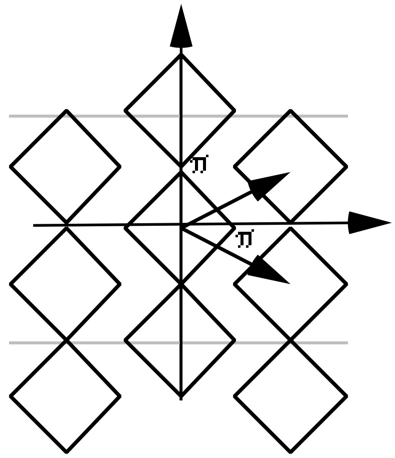

Figure 12

Since $|\det \mathbf{V}_1| = |\det \mathbf{V}_2| = 1$ the sampling density is the same in both cases. Notice that using rectangular sampling the sampling rate cannot be reduced without introducing aliasing. The sampling rate can be reduced in the hexagonally sampled signal by a factor of two in the x direction without introducing aliasing.

If the original signal is bandlimited then eq. 11 simplifies to

$$\mathbf{Y}(\omega) = \frac{1}{|\det \mathbf{V}|}\mathbf{X}(\mathbf{V}^{\mathrm{T}\,-1}\omega) \qquad (17)$$

for values of the spectrum between $-\pi$ and π. In this case the original signal can be recovered from the samples to give

$$\mathbf{x}(\mathbf{t}) = \frac{|\det \mathbf{V}|}{(2\pi)^{\mathbf{n}}}\sum_{\mathbf{n}} \mathbf{y}(\mathbf{n})\int_{B} \exp\left[j\omega^{\mathrm{T}}(\mathbf{t} - \mathbf{V}\mathbf{n})\right] d\Omega \qquad (18)$$

where the hypervolume of integration **B** is from $-\pi$ to π in each dimension of the signal.

References

FAHLE 1981 Fahle, M. and T. Poggio. “Visual Hyperacuity: Spatiotemporal Interpolation in Human Vision”. *Proc. R. Soc. Lond*., B 213, pp. 451-477, 1981.

MONTALVO 1979 Montalvo, Fanya S., chair of Special Session on Vision. *Computer Graphics*, *13*, 2, August 1979, pp. 121-125.

MEYER 1980 Meyer, Gary W. and Donald P. Greenberg. “Perceptual Color Spaces for Computer Graphics”. *Computer Graphics,14,* 3, 1980, pp. 254-261.

JOBLOVE 1978 Joblove, G.H. and Donald P. Greenberg. “Color Spaces for Computer Graphics”. *Computer Graphics*, August 1978.

FERWERDA 1988 Ferwerda, James A. and Donald P. Greenberg. “A psychophysical approach to assessing the quality of antialiased images”. *IEEE Computer Graphics and Applications*, V8, P85(11), Sept., 1988.

MITCHELL 1988 Mitchell, Dan P. and Arun N. Netravali. “Reconstruction Filters in Computer Graphics”. *Computer Graphics, 22,* 4, August 1988, pp. 221-228.

SCHREIBER 1985 Schreiber,William F. and Donald E. Troxel. “Transformation Between Continuous and Discrete Representations of Images: A Perceptual Approach”. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. PAMI-7, No. 2, March 1985, pp. 178-186.

DUBOIS 1988 Eric DuBois and William F. Schreiber. “Improvements to NTSC by Multidimensional Filtering”. *SMPTE Journal*, June 1988, pp. 446-463.

FEIBUSH 1980 Feibush, Eliot, Marc Levoy and Rob Cook. “Synthetic Texturing Using Digital Filtering”. *Computer Graphics*, *14*, 3, July 1980, pp. 294-301.

CATMULL 1984 Catmull, Edwin. “An Analytic Visible Surface Algorithm for Independent Pixel Processing”. *Computer Graphics*, *18*, 3, July 1984, pp. 109-106.

GRANT 1985 Grant, Charles W. “Integrated Analytic Spatial and Temporal Antialiasing”. ACM Computer Graphics, vol. 19, no. 3, July 1985, SIGGRAPH 1985 Proceedings, pp. 79-84.

LOBB 1987 Lobb, Richard J. “Antialiasing of Polygons with a Weighted Filter”. Computer Graphics 1987, Proceedings of CG International '87, Tosiyasu L. Kunii ed. pp.107-128 Springer Verlag Tokyo 1987

TURKOWSKI 1982 Turkowski, K. “Anti-aliasing through the use of coordinate transforms”. *ACM Transactions on Graphics,1*, (3), 1982, pp. 215-234.

ABRAM 1985	Abram, Greg, Lee Westover, and Turner Whitted. “Efficent Alias-Free Rendering Using Bit-Masks ans Look-Up Tables”. *Computer Graphics*, *19*, 3, July 1985, pp. 53-60.
GUENTER	Guenter, Brian. “The Psychophysics of Anisotropic Prefiltering”. In preparation.
OPPENHEIM 1975	Oppenheim, Alan. V., Schafer, Ronald W. “Digital Signal Processing”. Prentice-Hall 1975
MERSEREAU 1983	Mersereau, Russel M. and Speake, Theresa C. “The Processing of Periodically Sampled Multidimensional Signals”. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-31, No. 1, February 1983, pp. 188-194.
MITCHELL 1988	Mitchell, Dan P. and Arun N. Netravali. “Reconstruction Filters in Computer Graphics”. *Computer Graphics*, *22*, 4, August 1988, pp. 221-228.
CROCHIERE 1983	Crochiere, Ronald E., and Lawrence R. Rabiner. Multirate Digital Signal Processing. Prentice-Hall Signal Processing Series. 1983.
MERSEREAU 1984	Mersereau, Russel M. and Dudgeon, Dan E. Multidimensional Digital Signal Processing. Prentice Hall 1984
PETERSON 1962	Peterson, D. P. and D. Middleton. “Sampling and Reconstruction of Wave-Number limited Functions in N-Dimensional Euclidean Spaces”. *Information and Control*, 5 (1962), pp. 279-323.
WENDLAND 1984	Wendland, Broder. “On Picture Quality of Some Televised Signal Processing Techniques”. *SMPTE Journal*, October 1984, pp. 915-922.
WENDLAND 1983	Wendland, Broder. “Extended Definition Television with High Picture Quality”. *SMPTE Journal*, October 1983, pp. 1028-1035
CASSELS 1971	J.W.S. Cassels. An Introduction to the Geometry of Numbers. Springer-Verlag 1971